



TITLE:

部分観測可能なマルコフ連鎖におけるジョブ・サーチについて (不確実性の下での意思決定の数理)

AUTHOR(S):

中井, 達

CITATION:

中井, 達. 部分観測可能なマルコフ連鎖におけるジョブ・サーチについて (不確実性の下での意思決定の数理). 数理解析研究所講究録 2003, 1306: 91-100

ISSUE DATE:

2003-02

URL:

<http://hdl.handle.net/2433/42810>

RIGHT:

部分観測可能なマルコフ連鎖におけるジョブ・サーチについて

中井 達 (Tōru Nakai)

九州大学大学院経済学研究院 (Faculty of Economics, Kyushu University)

1 はじめに

ここでは、部分観測可能なマルコフ連鎖において、それぞれの状態に推移する確率の性質と、ジョブ・サーチにおいて最適政策にしたがって最適に振る舞ったときの同様の確率の性質を考える。このジョブ・サーチは、最適停止問題の1つであり、Lippman and MacCall [7] によって解析された動的な経済 (dynamic economy) の場合を拡張したものである。これは、それぞれの状態が経済の状態を表し、それぞれの状態に依存した、ジョブの大きさを表す確率変数を考える。[7] と異なる点は、部分観測可能なマルコフ連鎖にしたがって変化する状態を直接知ることができないことであり、この場合にそれぞれの状態に推移する確率の性質をみる。ジョブ・サーチにおいては、最適政策は一般的にあらかじめ決められた大きさ (reservation wage) によって定まるが、ここで考えるような不完備情報の場合には必ずしも当てはまらない。

いま、未知の状態に関する情報は、状態空間上の確率分布で表されるとし、学習プロセスとしてベイズの定理を用いる。そのため、total positive of order two または、簡単に TP_2 と呼ばれる性質を用いるが、これは尤度比を用いた順序関係と関連し、事前情報、事後情報、最適政策やそれぞれの状態に推移する確率との関係を考える。この TP_2 の性質については、Karlin and McGregor [2], Karlin [1] や Karlin and Rinott [3] などに詳しい。

部分観測可能なマルコフ連鎖におけるジョブ・サーチを考えるために、動的な経済におけるジョブ・サーチで、状態を直接知ることができる場合について基本的な性質を2節でまとめ、それぞれの状態に推移する確率が TP_2 であることを示す。3節では、状態が部分観測可能なマルコフ連鎖にしたがって推移する場合に、それぞれの状態に推移する確率が TP_2 であることが示される。ジョブ・サーチにおける最適政策にしたがった場合についても考える。つぎに、学習プロセスがベイズの定理にしたがった場合に、それぞれの状態に推移する確率の性質をみるために、状態を直接知ることができる場合についても考える。とくに、有限状態マルコフ連鎖で状態空間を $\{1, 2, \dots, K\}$ とするとき、状態 K が債務不履行 (default) の状態を表すとすれば、破産する確率の性質を求めることになっている。したがって、状態に関する情報が不完備な場合に、これらの確率の性質を求めることになっている。

2 動的な経済 (dynamic economy) におけるジョブ・サーチ

2.1 最適政策と期待利得

状態空間を $\{1, 2, \dots, K\}$ とし、推移行列を $P = (p_{ij})_{i,j=1,2,\dots,K}$ とするマルコフ連鎖を考える。このマルコフ連鎖の、それぞれの状態 i が経済の状態に対応するものと考え、これらの状態 ($i = 1, 2, \dots, K$) に対して、ジョブの大きさを表す確率変数 X_i を対応させる。また、つぎのジョブの大きさを知るためには、費用 c が必要である。このとき、割引率を β としたときの期待利得を最大にする最適政策と総期待利得を Lippman and MacCall [7] では、つ

ぎの仮定の下で議論している。

- (1) $F_i(x)$ は確率的に減少、すなわち、任意の x に対して $F_1(x) \geq F_2(x) \geq \dots \geq F_n(x)$ である。
- (2) 任意の k に対して、 $\sum_{j=k}^S p_{ij}$ は i に関して非減少である。

例えば、状態空間を $S = \{1, 2, \dots, K\}$ とするとき、推移確率 P において $p_{KK} = 1$ とし、状態 K のとき確率 1 で $X_K = 0$ とすれば、この状態 K は債務不履行 (default) の状態と考えることが出来る。すなわち、企業の状態を状態 k とし、確率変数 X_k を企業の状態が k のときのその企業の株価あるいは債券価格と考える。また、効用 $\frac{1-\delta^n}{1-\delta}x$ は、持っている株あるいは債権等を売却して安全資産へ投資することによって得られる利得と考えることが出来る。このとき、企業の状態と、その株価あるいは債券価格 x を知って売却するかどうかを決定することになる。

ここでは、動的な経済におけるジョブ・サーチをこのマルコフ連鎖の状態に関して不完備情報の場合に解析することを目的とするので、上記の場合とは異なり推移行列と確率変数 X_i ($i = 1, 2, \dots, K$) の分布関数に関して、つぎの2つの仮定 (仮定1と2) のもとで議論する。状態空間が $\{1, 2, \dots, K\}$ で、推移行列が $P = (p_{ij})_{i,j=1,2,\dots,K}$ のマルコフ連鎖において、それぞれの状態 i に対応する確率変数 X_i は絶対連続で密度関数 $f_i(x)$ ($i = 1, 2, 3, \dots, K$) を持つものとする。ここでの議論は、Nakai [11] のように一般の部分観測可能なマルコフ過程の場合にも拡張することは可能であり、多段決定過程に応用することもできる (Nakai[9, 10]etc.)。まずはじめに、定義1によって尤度比を用いて、確率変数のあいだに順序関係を導入する。

定義1 確率密度関数 $f(x), g(x)$ をもつ2つの確率変数 X, Y に対して、任意の x と y が $x \geq y$ であれば、 $f(y)g(x) \leq f(x)g(y)$ となるとき、 X は Y より尤度比の意味で大きいといい、簡単に $X \succeq Y$ と表す。

定義2 $K \times K$ 行列 $A = (a_{ij})_{i,j=1,2,3,\dots,K}$ に対して、 $a_{ik}a_{jl} \geq a_{jk}a_{il}$, i.e., $\begin{vmatrix} a_{ik} & a_{il} \\ a_{jk} & a_{jl} \end{vmatrix} \geq 0$ が、 $i \leq j$ および $k \leq l$ を満たす任意の i, j, k と l に対して成り立つとき ($i, j = 1, 2, \dots, K$, $k, l = 1, 2, \dots, K$)、この行列 A は *total positive of order two* といい、簡単に TP_2 と表す。

このとき、つぎの性質が成り立つ。

補題1 行列 A と B が TP_2 ならば、積 AB もまた TP_2 である。

動的な経済におけるジョブ・サーチを不完備情報の場合に考えるために、推移行列 P と確率変数 X_i ($i = 1, 2, 3, \dots, K$) に対して、2つの仮定 (仮定1と2) を導入する。仮定1と2は、部分観測可能なマルコフ連鎖において、ベイズの定理を学習プロセスとして用いることから必要となるものである。

仮定1 確率変数 $\{X_i \mid i = 1, 2, \dots, K\}$ に対して、 $k \leq l$ ならば $X_k \succeq X_l$ である ($k, l = 1, 2, \dots, K$)。言い換えれば、 X_i は i に関して尤度比の意味で減少する。

仮定2 推移行列 $P = (p_{ij})_{i,j=1,2,3,\dots,K}$ は TP_2 である。

仮定1において、 $X_k \succeq X_l$ だから、もし $x > y$ であれば、任意の k と l ($k \leq l, k, l = 1, 2, \dots, K$) に対して $f_k(y)f_l(x) \leq f_k(x)f_l(y)$ となっている。言い換えれば、 $x > y$ のとき、 $k \leq l$ ($k, l = 1, 2, \dots, K$) ならば $f_l(x)f_k(y) - f_k(x)f_l(y) \leq 0$ である。仮定1から、確率変数

X_i は、 i が大きくなれば、小さな値を取りやすくなる。すなわち、状態 1 が最も良い状態であり、状態 2 は 2 番目に良く、 \dots 、状態 K は最も悪い状態である。仮定 2 は、マルコフ連鎖において TP_2 として知られている。この仮定より、状態の数 i 大きくなればなるほど、悪い状態に入る確率が増加する。

つぎに、 $u_n(x)$ を、残り計画期間が n で、直面している仕事の大きさが x のときの利得とする。このとき、つぎの仕事を探すための費用を c とし、割引率を $0 < \beta < 1$ とする。このとき、 $v_n(i, x)$ を残り計画期間が n で、直面している仕事の大きさが x のとき、最適に振舞って得られる割引率 β で割引かれた総期待利得とすれば、最適性の原理から、 $v_n(i, x)$ はつぎの再帰方程式を満たす。

$$v_{n+1}(i, x) = \max \left\{ u_n(x), -c + \beta \sum_{j=1}^K p_{ij} \int_0^\infty v_n(j, y) dF_j(y) \right\} \quad (1)$$

ただし、 $v_1(i, x) = u_1(x)$ とする。ここで、 $u_n(x)$ は x と n に関する増加関数とする。例えば、 $u_n(x) = \frac{1-\delta^n}{1-\delta}x$ は条件を満足する。Lippman and MacCall [7] と同じように、計画期間が n でマルコフ連鎖の状態が i のとき、最適政策は reservation wage $\alpha_n(i)$ によって定まり、帰納法から $\alpha_n(i)$ と $v_n(i, x)$ はつぎの性質を持つ。

補題 2 任意の状態 $i \in \{1, 2, \dots, K-1\}$ と正整数 n に対し、 $\alpha_n(i) \geq \alpha_n(i+1)$ である。また、任意の状態 $i \in \{1, 2, \dots, K-1\}$ と正整数 n に対して、 $\alpha_{n+1}(i) \geq \alpha_n(i)$ である。

補題 3 任意の状態 $i \in \{1, 2, \dots, K-1\}$ と正整数 n に対し $v_{n+1}(i, x) \geq v_n(i, x)$, $v_n(i, x) \geq v_n(i+1, x)$ である ($x > 0$)。もし、 $x > y$ ならば $v_n(i, x) \geq v_n(i, y)$ である。

これらの補題は、仮定 1・2 と $u_n(x)$ に関する仮定から、 n に関する帰納法を用いて示すことが出来る。また、 P が仮定 2 を満たすから、 $k \leq l$ となる任意の k と l に対して $(k, l = 1, 2, \dots, K)$ 、 $\begin{vmatrix} p_{ik} & p_{il} \\ p_{i+1k} & p_{i+1l} \end{vmatrix} \geq 0$ となる。したがって、 $p_i = (p_{i1}, \dots, p_{iK})$ および $p_{i+1} = (p_{i+11}, \dots, p_{i+1K})$ とおけば、任意の $i \in \{1, 2, \dots, K\}$ に対して $p_{i+1} \succeq p_i$ となる。よって、 $p_{i+1} \succeq p_i$ と系 1 より $\sum_{j=1}^K p_{ij} \int_0^\infty v_n(j, y) dF_j(y) = \int_0^\infty \sum_{j=1}^K p_{ij} v_n(j, y) f_j(y) dy$ だから、 $\sum_{j=1}^K p_{ij} \int_0^\infty v_n(j, y) dF_j(y) \geq \sum_{j=1}^K p_{i+1j} \int_0^\infty v_n(j, y) dF_j(y)$ となる。このとき、つぎの性質が成り立つ。

補題 4 j に関する減少列 $\{a_j\}_{j=1, \dots, K}$ で、 $\sum_1^K a_j = 0$ であり、適当な $k (1 < k < K)$ に対して $a_j \begin{cases} \geq \\ \leq \end{cases} 0$ if $j \begin{cases} \leq \\ > \end{cases} k$ とする。いま、 $h(i, x)$ を i に関して減少し、 x に関して増加する関数とすれば $\int_0^\infty \sum_{j=1}^K a_j h(j, y) f_j(y) dy \geq 0$ である。

ここで、補題 4 から、つぎの系 1 が導かれる。

系 1 $\mu \succeq \nu$ ならば、 $\int_0^\infty \sum_{j=1}^K h(j, x) \nu_j f_j(x) dx \geq \int_0^\infty \sum_{j=1}^K h(j, x) \mu_j f_j(x) dx$ となる ($\mu, \nu \in \mathcal{S}$)。

2.2 状態への推移確率

マルコフ連鎖の状態を知ることができたとき、 n 期間後に状態 j になる確率を考える。そのため、はじめにマルコフ連鎖の状態の推移だけに着目する。マルコフ連鎖の状態が i のとき、 $\bar{p}_{ij}(n)$ を n 期間後に状態 j となる確率とする ($i, j = 1, 2, \dots, K, n = 1, 2, \dots$)。このとき、確率 $\bar{p}_{ij}(n)$ は、 $\bar{p}_{ij}(1) = p_{ij}$ を初期値とする、再帰方程式 $\bar{p}_{ij}(n) = \sum_{k=1}^K p_{ik} \bar{p}_{kj}(n-1)$ を満たす。ここで、 $K \times K$ 行列を $\bar{P}(n) = (\bar{p}_{ij}(n))$ とすれば、 $\bar{P}(1) = P$ および $\bar{P}(n) = P \bar{P}(n-1)$ と表せる。また、仮定 2 より $P = (p_{ij})$ が TP_2 だから、 $\bar{P}(n-1)$ が TP_2 であれば、 n に関する帰納法によって、補題 1 より $\bar{P}(n) = (\bar{p}_{ij}(n))$ は TP_2 となる。

ここでは、状態の変化が推移確率を $\bar{p}_{ij}(n)$ とするマルコフ連鎖にしたがう場合であったが、動的な経済におけるジョブ・サーチの場合に考える。マルコフ連鎖の状態が i で、計画期間が n のときに、ジョブ・サーチの最適政策にしたがったとき、 $\bar{p}_{ij}(n, m)$ を m 期間後の状態が j となる確率とする ($i, j = 1, 2, \dots, K, n, m = 1, 2, \dots, m \leq n$)。ジョブ・サーチにおいて、マルコフ連鎖の状態を直接知ることができる場合には、最適政策は reservation wages $\alpha(i, n)$ によって定まった。ここで、 $F_i(\alpha(i, n))$ は、マルコフ連鎖の状態が i で n 期間残されているとき、直面しているジョブを採用しない確率である。したがって、 $\bar{p}_{ij}(n, m)$ が、つぎの再帰方程式を満たすことは簡単にわかる。

$$\bar{p}_{ij}(n, m) = F_i(\alpha(i, n)) \sum_{k=1}^K p_{ik} \bar{p}_{kj}(n+1, m-1) \quad (2)$$

ここで、初期条件は $\bar{p}_{ij}(n, 1) = F_i(\alpha(i, n)) p_{ij}$ である。いま、 $\bar{P}(n, m) = (\bar{p}_{ij}(n, m))$ とおけば、任意の $n (> 0)$ に対して $\bar{P}(n, 1) = D(F_1(\alpha(1, n)), \dots, F_K(\alpha(K, n))) P$ であり、(2) 式から

$$\bar{P}(n, m) = D(F_1(\alpha(1, n)), \dots, F_K(\alpha(K, n))) P \bar{P}(n+1, m-1), \quad (3)$$

となる。ただし、 $D(F_1(\alpha(1, n)), \dots, F_K(\alpha(K, n)))$ はつぎのような対角行列である。

$$D(F_1(\alpha(1, n)), \dots, F_K(\alpha(K, n))) = \begin{pmatrix} F_1(\alpha(1, n)) & 0 & \dots & 0 \\ 0 & F_2(\alpha(2, n)) & \dots & 0 \\ \vdots & \vdots & \dots & \vdots \\ 0 & 0 & \dots & F_K(\alpha(K, n)) \end{pmatrix}$$

系 2 行列 A が TP_2 で、 D を対角行列とすれば、積 DA もまた TP_2 である。

このとき、これらの $\bar{P}(n, m)$ は、つぎの補題を満たす。

命題 1 $\bar{P}(n, m) = (\bar{p}_{ij}(n, m))$ は TP_2 である。

証明: m に関する帰納法を用いる。 $m = 1$ のとき、補題 2 より

$$\bar{P}(n, 1) = D(F_1(\alpha(1, n)), \dots, F_K(\alpha(K, n))) P$$

は TP_2 である。 m 以下の値に対して $\bar{P}(n, m)$ が TP_2 であれば、 $P = (p_{ij})$ と $\bar{P}(n+1, m-1)$ が TP_2 だから、補題 2 より $\bar{P}(n, m) = (\bar{p}_{ij}(n, m))$ もまた TP_2 である。□

3 不完備情報のジョブ・サーチ

3.1 最適政策と期待利得

動的な経済でのジョブ・サーチで、マルコフ連鎖の状態を直接知ることができない場合、すなわち部分観測可能なマルコフ連鎖にしたがう場合を考える。ここで、状態空間を $S = \{1, 2, \dots, K\}$ とするとき、推移確率 P において $p_{KK} = 1$ とし、状態 K のとき確率 1 で $X_K = 0$ とすれば、この状態 K は債務不履行 (default) の状態と考えることが出来る。すなわち、企業の状態とその格付けの関係を、状態 k に依存する確率変数 $X_k = 0$ で表せば、その企業の格付けを情報プロセスと考えて、企業の状態を類推することになる。また、効用 $u_n(x)$ は、企業の格付けが x のとき、持っている株あるいは債権等を売却して安全資産へ投資することによって得られる利得と考えればよい。

このマルコフ連鎖の状態についての情報は、状態空間 $\{1, 2, \dots, K\}$ 上の確率分布 μ で表されているとし、 S を未知の状態に関する情報全体の集合とすれば

$$S = \left\{ \mu = (\mu_1, \mu_2, \dots, \mu_K) \mid \sum_{i=1}^K \mu_i = 1, \mu_i \geq 0 (i = 1, 2, \dots, K) \right\}$$

となる。

つぎに、 S に含まれる情報のあいだに、定義 3 にしたがって尤度比を用いて順序関係を仮定する。この確率的な順序関係は一般化することができ Nakai[11, 12] に詳しい。また、多段決定問題に対する応用も考えられている。

定義 3 状態空間 $\{1, 2, \dots, K\}$ 上の 2 つの確率分布 μ, ν に対して、任意の $i, j (i \leq j, i, j = 1, 2, \dots, K)$ に対して、 $\mu_j \nu_i \geq \mu_i \nu_j$ であり、少なくとも 1 つの i と j の組み合わせに対して不等号が成り立つとき、 μ は尤度比の意味で ν より大きいといい、簡単に $\mu \succ \nu$ と表す。

この順序は半順序であり、total positive of order two または、簡単に TP_2 と呼ばれる。定義 3 より、 $\mu \succeq \nu$ ならば $(\mu, \nu \in S)$ 、任意の k と $l (k \leq l, k, l = 1, 2, \dots, K)$ に対して $\mu_l \nu_k \geq \mu_k \nu_l$ だから、 j が大きくなるにしたがって、状態が j である確率の比 $\frac{\nu_j}{\nu_i}$ は増加する。

つぎに、未知の状態に関する情報を得るプロセスとして、状態に依存するジョブの大きさを表す確率変数 $\{X_i \mid i = 1, 2, 3, \dots, K\}$ を考える。すなわち、現れたジョブの大きさから未知の状態に関する情報を改良する。状態に関する情報が μ のとき、新たに現れたジョブの大きさが x であれば、ベイズの定理を用いて未知の状態についての情報を $\mu(x) = (\mu(x)_1, \mu(x)_2, \dots, \mu(x)_K)$ と改良する。このあと、推移行列 P にしたがって状態が推移し、推移した後の状態に関するつぎの時点での事前情報を $\overline{\mu(x)}$ とする。

ここで、ベクトル値関数 $h(x)$ に関して、定義 4 によって単調性を定義する。

定義 4 x のベクトル値関数 $h(x) = (h_1(x), h_2(x), \dots, h_K(x))$ が、 $x > y$ のとき $h(y) \succeq h(x)$ である ($h(x) \succeq h(y)$)、すなわち $h_j(y) h_i(x) \geq h_i(y) h_j(x)$ ($h_j(y) h_i(x) \leq h_i(y) h_j(x)$) が任意の j と $i (i \leq j, i, j = 1, 2, \dots, K)$ に対して成り立つとする。このとき、この関数 $h(x)$ を、 x に関する減少 (増加) 関数という。

密度関数を $\{f_i(x) \mid i = 1, 2, \dots, K\}$ とする確率変数 $\{X_i \mid i = 1, 2, \dots, K\}$ が仮定 1 を満たすから、 $f(x) = (f_1(x), \dots, f_K(x))$ は $f(y) \succeq f(x)$ を満たす。すなわち、 $x > y$ ならば、任意の $i, j (i \leq j, i, j = 1, 2, \dots, K)$ に対して $f_i(y) f_j(x) \leq f_i(x) f_j(y)$ となる。したがって、 $f(x)$ は x の増加関数である。

事前情報 μ と事後情報 $\overline{\mu(x)}$ のあいだに、仮定 1 と 2 のもとでつぎの性質が成り立つ。これは、Nakai[11] の補題 5 など知られている。

補題 5 $\mu \succ \nu$ ならば $(\mu, \nu \in S)$ 、任意の x に対して $\mu(x) \succ \nu(x)$ および $\overline{\mu(x)} \succ \overline{\nu(x)}$ である。また、 $\mu \succ \nu$ ならば $(\mu, \nu \in S)$ 、 $\mu(x)$ と $\overline{\mu(x)}$ は x の減少関数である。

仮定 1 と 2 のもとで、補題 5 より、事前情報 μ のあいだでの定義 3 による順序関係は、事後情報 $\mu(x)$ およびつぎの期での $\overline{\mu(x)}$ においても保存される。さらに、同じ事前情報 μ であっても、現れたジョブの大きさが大きくなれば、事後情報 $\overline{\mu(x)}$ も定義 3 の意味で良くなる。

状態についての事前情報が μ のジョブ・サーチに対して、 $v_n(\mu, x)$ を計画期間が n で、利用可能なジョブの大きさが x のとき ($0 \leq \beta \leq 1$)、割引率を β としたときの最適政策にしたがったときの期待利得とすれば、最適性の原理から、 $v_n(\mu, x)$ は再帰方程式

$$v_n(\mu, x) = \max \left\{ u_n(x), c + \beta \int_0^\infty v_n(\overline{\mu(x)}, y) dF_{\overline{\mu(x)}}(y) \right\}. \quad (4)$$

を満たす。つぎに $S(\mu, n) = \left\{ x \mid u_n(x) \geq c + \beta \int_0^\infty v_n(\overline{\mu(x)}, y) dF_{\overline{\mu(x)}}(y) \right\}$ とし、 $C(\mu, n) = S(\mu, n)^c$ とおけば、集合 $S(\mu, n)$ と $C(\mu, n)$ はそれぞれこのジョブ・サーチに対する停止領域と継続領域に対応する。つぎの補題は、Nakai[8] の補題 3.4 から導かれるが、Nakai[11, 9, 8] と同じように一般化することも可能である。

補題 6 $h(x)$ が x の非減少関数のとき、 $\mu \succeq \nu$ ならば、 $\int_0^\infty h(x) dF_\mu(x) \leq \int_0^\infty h(x) dF_\nu(x)$ である ($\mu, \nu \in S$)。

また、被積分関数 $v_n(\overline{\mu(x)}, z)$ が z の増加関数ならば、補題 6 より、 $x > y$ ならば

$$\int_0^\infty v_n(\overline{\mu(x)}, z) dF_{\overline{\mu(x)}}(z) \geq \int_0^\infty v_n(\overline{\mu(y)}, z) dF_{\overline{\mu(y)}}(z)$$

である。したがって、これらの領域 $S(\mu, n)$ と $C(\mu, n)$ に対して、(4) 式からつぎの性質が得られる。

補題 7 $\mu \succeq \nu$ ($\mu, \nu \in S$) ならば、 $S(\nu, n) \subset S(\mu, n)$ および、 $S(\mu, n+1) \subset S(\mu, n)$ である。

$S(\mu, n) \cup C(\mu, n) = \mathcal{R}_+$ であり、任意の μ と $n \geq 1$ に対して $S(\mu, n) \cap C(\mu, n) = \emptyset$ だから、この補題より $C(\mu, n) \subset C(\nu, n)$ および $C(\mu, n) \subset C(\mu, n+1)$ となる。値 $v_n(\mu, x)$ もまたつぎの性質を持つ。

補題 8 $\mu \succeq \nu$ ならば $(\mu, \nu \in S)$ 、 $v_n(\mu, x) \leq v_n(\nu, x)$ である。 $x > y$ のとき $v_{n+1}(\mu, x) \geq v_n(\mu, x)$ および $v_n(\mu, x) \geq v_n(\mu, y)$ となる。

これらの性質は、Nakai [11] などと同様に n に関する帰納法で示すことが出来る。

3.2 状態へ推移する確率—不完備情報の場合

2.2 節と同様に、状態が部分観測可能なマルコフ連鎖にしたがって推移するとき、仮定 1 と 2 のもとで、 n 期間後に状態 j に推移する確率を考える。

はじめに、ジョブ・サーチにおける決定とジョブの大きさを用いた学習を除いて、これらの確率を考える。状態についての事前情報が μ のとき、 $\bar{P}(\mu, m)$ を m 期間後に状態が j となる確率とし、初期条件として $m = 1$ のとき $\bar{P}(\mu, 1) = \left(\sum_{i=1}^K \mu_i p_{i1}, \dots, \sum_{i=1}^K \mu_i p_{iK} \right) = \mu P$ とする。このことから、 $m = 2$ に対して $\bar{P}(\mu, 2) = \bar{P}(\bar{\mu}, 1) = \bar{\mu} P = \mu P^2$ であり、 $\bar{P}(\mu, m)$ の再帰方程式はつぎのようになる。

$$\bar{P}(\mu, m) = \bar{P}(\bar{\mu}, m-1) = \bar{P}(\mu P, m-1) = \mu P^m. \quad (5)$$

また、 P が TP_2 だから、 m に関する帰納法より $\bar{P}(\mu, m) = \mu P^m$ もまた TP_2 である。

補題 9 $\mu \succeq \nu$ であり $(\mu, \nu \in \mathcal{S})$ 、 $A = (a_{ij})$ が TP_2 ならば $(\mu, \nu \in \mathcal{S})$ 、 $\mu A \succeq \nu A$ である。

したがって、(5) 式を用いて、補題 9 からつぎの命題が導かれる。

命題 2 $\mu \succeq \nu$ ならば $(\mu, \nu \in \mathcal{S})$ 、 $\bar{P}(\mu, m) \succeq \bar{P}(\nu, m)$ である。

つぎに、決定を除いてこれらの確率を考える。すなわち、ジョブの大きさを用いて未知の状態に関する学習を考える。ここで、推移と学習の関係は、はじめにジョブの大きさを状態についての情報として観測し、状態についての情報を改良する。そのあと、推移行列 P にしたがって新しい状態へ推移すると考える。したがって、状態についての情報が μ のとき、推移がすでに終わっていることになる。状態についての事前情報が μ のとき、 $\hat{P}(\mu, m)_j$ を m 期間後に状態が j となる確率とし ($j = 1, 2, \dots, K$)、 $\hat{P}(\mu, m) = (\hat{P}(\mu, m)_1, \hat{P}(\mu, m)_2, \dots, \hat{P}(\mu, m)_K)$ とおく。

ここで、ベクトル値関数 $u(x) = (u_1(x), \dots, u_K(x))$ に対して、 $\int_a^b u_i(x) dF(x)$ が任意の i について存在するとき ($i = 1, 2, \dots, K$)、

$$\int_a^b u(x) dF(x) = \left(\int_a^b u_1(x) dF(x), \dots, \int_a^b u_K(x) dF(x) \right)$$

を $\int_a^b u(x) dF(x)$ と表す。

事前情報が μ のとき、 $\hat{P}(\mu, 1) = (\hat{P}(\mu, 1)_1, \dots, \hat{P}(\mu, 1)_K)$ は、学習と推移のあとでの状態空間上の確率分布だから、

$$\hat{P}(\mu, 1) = \int_0^\infty \mu(x) P dF_\mu(x) = \int_0^\infty \overline{\mu(x)} dF_\mu(x)$$

となる。 n 期での事前情報が μ のとき、 $\hat{P}(\mu, 2)$ は 2 期あとでの状態空間上の確率分布であり、学習のあとでの $(n+1)$ 期における情報が $\overline{\mu(x)}$ だから、 $\hat{P}(\mu, 2) = \int_0^\infty \hat{P}(\overline{\mu(x)}, 1) dF_\mu(x)$ となる。

同じように、未知の状態に関する事前情報が μ であれば、 $\hat{P}(\mu, m)$ は m 期後の状態空間上の確率分布だから、この $\hat{P}(\mu, m)$ は、つぎの再帰方程式を満たす。

$$\hat{P}(\mu, m) = \int_0^\infty \hat{P}(\overline{\mu(x)}, m-1) dF_\mu(x) \quad (6)$$

ここで、 $\hat{P}(\mu, 1) = \int_0^\infty \overline{\mu(x)} dF_\mu(x)$ とする。これらの確率分布の性質を求めるために、定義 5 によって順序を定義する。

定義 5 2つの x の関数 $g(x) = (g_1(x), \dots, g_K(x))$ と $h(x) = (h_1(x), \dots, h_K(x))$ に対して、 $g(x)_j h(x)_i \geq g(x)_i h(x)_j$ が任意の i と j に対して成り立つとき ($i \leq j, i, j = 1, 2, \dots, K$)、 $g(x)$ は $h(x)$ より TP_2 の意味で大きいといい、 $g(x) \succeq h(x)$ と表す。

補題 10 $g(x) = (g(x)_1, \dots, g(x)_K)$ と $h(x) = (h(x)_1, \dots, h(x)_K)$ が x に関して減少関数であり、 $g(x) \succeq h(x)$ ならば $\int_0^\infty g(x) dF(x) \succeq \int_0^\infty h(x) dF(x)$ である。

系 3 $\mu \succeq \nu$ ($\mu, \nu \in S$) ならば、 $\int_0^\infty \overline{\mu(x)} dF(x) \succeq \int_0^\infty \overline{\nu(x)} dF(x)$ である。

補題 11 $\mu \succeq \nu$ であり ($\mu, \nu \in S$)、ベクトル値関数 $h(x)$ が x に関して減少関数ならば、仮定 1 と 2 のもとで $\int_0^\infty h(x) dF_\mu(x) \succeq \int_0^\infty h(x) dF_\nu(x)$ である。

もし、 $\mu \succeq \nu$ ならば ($\mu, \nu \in S$)、補題 5 より $\overline{\mu} \succeq \overline{\nu}$ および $\overline{\mu(x)} \succeq \overline{\nu(x)}$ であり、したがって $\hat{P}(\mu, m)$ はつぎの性質を持つ。

$\overline{\mu(x)} \succeq \overline{\nu(x)}$ より、 $\hat{P}(\overline{\mu(x)}, 1) \succeq \hat{P}(\overline{\nu(x)}, 1)$ である。

$$\int_0^\infty g(x) dF_\mu(x) \succeq \int_0^\infty g(x) dF_\nu(x) \succeq \int_0^\infty h(x) dF_\nu(x)$$

より、補題 10 と 11 から、つぎの系が導かれる。

系 4 $\mu \succeq \nu$ のとき $g(x) \succeq h(x)$ ならば、 $\int_0^\infty g(x) dF_\mu(x) \succeq \int_0^\infty h(x) dF_\nu(x)$ である ($\mu, \nu \in S$)。

系 5 $\mu \succeq \nu$ ならば ($\mu, \nu \in S$)、仮定 1 と 2 のもとで、

$$\int_0^\infty h(\mu, x) dF_\mu(x) = \left(\int_0^\infty \mu(x)_1 dF_\mu(x), \dots, \int_0^\infty \mu(x)_K dF_\mu(x) \right) \succeq \int_0^\infty h(\nu, x) dF_\nu(x)$$

である。ただし、 $h(\mu, x)$ は μ の増加関数とし、 x の非増加関数とする。

命題 3 $\mu \succeq \nu$ ならば ($\mu, \nu \in S$)、 $\hat{P}(\mu, m)$ は x の増加関数である。すなわち、 $\hat{P}(\mu, m) \succeq \hat{P}(\nu, m)$ である。

証明: m に関する帰納法を用いる。 $m = 1$ のとき、 $\mu \succeq \nu$ ならば、 $\hat{P}(\mu, 1) = \int_0^\infty \overline{\mu(x)} dF_\mu(x)$ と系 3 より $\hat{P}(\mu, 1) \succeq \hat{P}(\nu, 1)$ となる。

系 5 から

$$\hat{P}(\mu, 2) = \int_0^\infty \hat{P}(\overline{\mu(x)}, 1) dF_\mu(x) \succeq \int_0^\infty \hat{P}(\overline{\nu(x)}, 1) dF_\mu(x) = \hat{P}(\nu, 2),$$

となり、このことから $\hat{P}(\mu, 2) \succeq \hat{P}(\nu, 2)$ である。

帰納法の仮定より、 $\mu \succeq \nu$ ならば $\hat{P}(\mu, m-1) \succeq \hat{P}(\nu, m-1)$ である。 $\overline{\mu(x)} \succeq \overline{\nu(x)}$ だから、 $\hat{P}(\overline{\mu(x)}, m-1) \succeq \hat{P}(\overline{\nu(x)}, m-1)$ も導かれる。系 5 より

$$\begin{aligned} \hat{P}(\mu, m) &= \int_0^\infty \hat{P}(\overline{\mu(x)}, m-1) dF_\mu(x) \\ &\succeq \int_0^\infty \hat{P}(\overline{\nu(x)}, m-1) dF_\mu(x) \\ &\succeq \int_0^\infty \hat{P}(\overline{\nu(x)}, m-1) dF_\nu(x) = \hat{P}(\nu, m), \end{aligned}$$

最後に、ジョブ・サーチにおいてジョブの大きさを用いて学習を行い、最適政策にしたがったときの確率を考える。計画期間が n のとき、状態に関する事前情報を μ とし、まだジョブの大きさを知らないとする。いま、 $\tilde{P}(\mu, n, m)_j$ を、最適政策にしたがったとき、 m 期あとに状態が j となる確率とする ($i, j = 1, 2, \dots, K, n, m = 1, 2, \dots, m \leq n$)。

はじめに、 $m = 1$ のときを考える。まず、推移と決定の順序関係をつぎのように考える。状態に関する情報としてジョブの大きさ x を知ってから、状態についての情報を改良し、このジョブを採択するかどうかを決定する。そのあと、推移行列 P にしたがって状態が推移し、新しい状態になる。

未知の状態についての学習と現れたジョブに対する決定が終了したとき、未知の状態についての事後情報を μ とおく。いま、 $\tilde{P}'(\mu, n, 1)_j$ を推移したあとでつぎの期に状態が j となる確率とすれば ($j = 1, 2, \dots, K$)、 $\tilde{P}'(\mu, n, 1) = (\tilde{P}'(\mu, n, 1)_1, \tilde{P}'(\mu, n, 1)_2, \dots, \tilde{P}'(\mu, n, 1)_K)$ は $\tilde{P}'(\mu, n, 1)_j = \sum_{i=1}^K \mu_i p_{i,j}$ および $\tilde{P}'(\mu, n, 1) = \mu P = \bar{\mu}$ を満たす。このことから、

$$\tilde{P}(\mu, n, 1)_j = \int_{C(\mu, n)} \tilde{P}'(\mu(x), n, 1)_j dF_{\mu}(x) = \int_{C(\mu, n)} \sum_{i=1}^K \mu(x)_i p_{i,j} dF_{\mu}(x)$$

となる。したがって、

$$\tilde{P}(\mu, n, 1) = \int_{C(\mu, n)} \tilde{P}'(\mu(x), n, 1) dF_{\mu}(x) = \int_{C(\mu, n)} \bar{\mu}(x) dF_{\mu}(x)$$

である。計画期間が n のとき未知の状態についての事前情報が μ のとき、 $\tilde{P}(\mu, n, m)_j$ を最適政策にしたがったとき m 期あとに状態が j となる確率とする ($i, j = 1, 2, \dots, K, n, m = 1, 2, \dots, m \leq n$)。この場合、未知の状態に依存した新しいジョブの大きさを知って、採択するかどうかを決定する。このジョブ・サーチは、 $x \in C(\mu, n)$ のときに、つぎの期へ進むから、 $\tilde{P}(\mu, n, m)_j$ がつぎの再帰方程式を満たすことは簡単にわかる。

$$\tilde{P}(\mu, n, m)_j = \int_{C(\mu, n)} \tilde{P}(\bar{\mu}(x), n+1, m-1)_j dF_{\mu}(x) \quad (7)$$

であり $\tilde{P}(\mu, n, m) = (\tilde{P}(\mu, n, m)_1, \dots, \tilde{P}(\mu, n, m)_K)$ である。 $\int_{S(\mu, n)} dF_{\mu}(x)$ がこの期でジョブを採択する確率だから、 $\sum_{j=1}^K \tilde{P}(\mu, n, m)_j \leq 1$ となる。さらに、 $\mu \succeq \nu$ より $C(\mu, n) \subset C(\nu, n)$ となる。すなわち、 μ が増加するにつれて、つぎの期に進む確率は減少し、その反対に悪い状態になる確率は増加する。マルコフ連鎖の状態を知ることができれば、命題 1 から $\bar{P}(n, m)$ は TP_2 であった。しかし、確率 $\tilde{P}(\bar{\mu}(x), n+1, m-1)_j$ が、観測した x によって変化するから、 $\tilde{P}(\mu, n, m)$ に関してそのような性質を求めることは難しい。

また、 TP_2 を用いて、いくつかの定義を行った (定義 1, 2, 3 と 4) い、 TP_2 (定義 1, 2, 3 と 4) の定義を用いて、仮定 1 と 2 のもとで得られた。 TP_2 については、Kijima[4] と Kijima and Ohnishi[5, 6] などでも TP_2 の性質を解析している。

参考文献

- [1] S. Karlin, *Total Positivity*, Stanford University Press, Stanford, California (1968).

- [2] S. Karlin and J. L. McGregor, Classical Diffusion Process and Total Positivity, *Journal of Mathematical Analysis and Applications*, **1**, 163–183 (1960).
- [3] S. Karlin and Y. Rinott, Total Positivity Properties of Absolute Value Multinomial Variables with Applications to Confidence Interval Estimates and Related Probabilistic Inequalities, *The Annals of Statistics*, **9**, 1035–1049 (1981).
- [4] M. Kijima, Monotonicities in a Markov Chain Model for Valuing Corporate Bonds Subject to Credit Risk, *Mathematical Finance*, **8**, 229–47 (1998).
- [5] M. Kijima and M. Ohnishi, Portfolio Selection Problems via the Bivariate Characterization of Stochastic Dominance Relations, *Mathematical Finance*, **6**, 237–277 (1996).
- [6] M. Kijima and M. Ohnishi, Stochastic Orders and Their Applications in Financial Optimization, *Mathematical Methods of Operations Research*, **50**, 351–372 (1999).
- [7] S. A. Lippman and J. J. McCall, Job Search in a Dynamic Economy, *Journal of Economic Theory*, **12**, 365–390 (1976).
- [8] T. Nakai, The Problem of Optimal Stopping in a Partially Observable Markov Chain, *Journal of Optimization Theory and Applications*, **45**, 425–442 (1985).
- [9] T. Nakai, A Sequential Stochastic Assignment Problem in a Partially Observable Markov Chain, *Mathematics of Operations Research*, **11**, 230–240 (1986).
- [10] T. Nakai, An Optimal Assignment Problem for Multiple Objects per Period – Case of a Partially Observable Markov Chain, *Bulletin of Informatics and Cybernetics*, **31**, 23–34 (1999).
- [11] T. Nakai, A Generalization of Multivariate Total Positivity of Order Two with an Application to Bayesian Learning Procedure, *Journal of Information & Optimization Sciences*, **23**, 163–176 (2002).
- [12] T. Nakai, A Generalized MTP₂ and a Sequential Stochastic Model on a Partially Observable Markov Process, *Probabilistic Methods in Discrete Mathematics—Proceedings of the Fifth International Petrozavodsk Conference*, (Eds. V.F. Kolchin, V.Ya. Kozlov, V.V. Mazalov, Yu.L. Pavlov and Yu.V. Prokhorov), VSP publishes, The Netherlands, 291–302 (2002).
- [13] T. Nakai, Properties of a Partially Observable Markov Chain for a Job Search Problem in a Dynamic Economy, *Proceedings of the 2nd Euro Japanese Workshop on Stochastic Modelling for Finance, Insurance, Production and Reliability*, (Eds. T. Dohi, N. Limonios and S. Osaki), Systems Reliability Engineering Laboratory, Hiroshima University, 340–349, 2002.
- [14] S. M. Ross, *Applied Probability Models with Optimization Applications*, Holden-Day, San Francisco, California (1970).